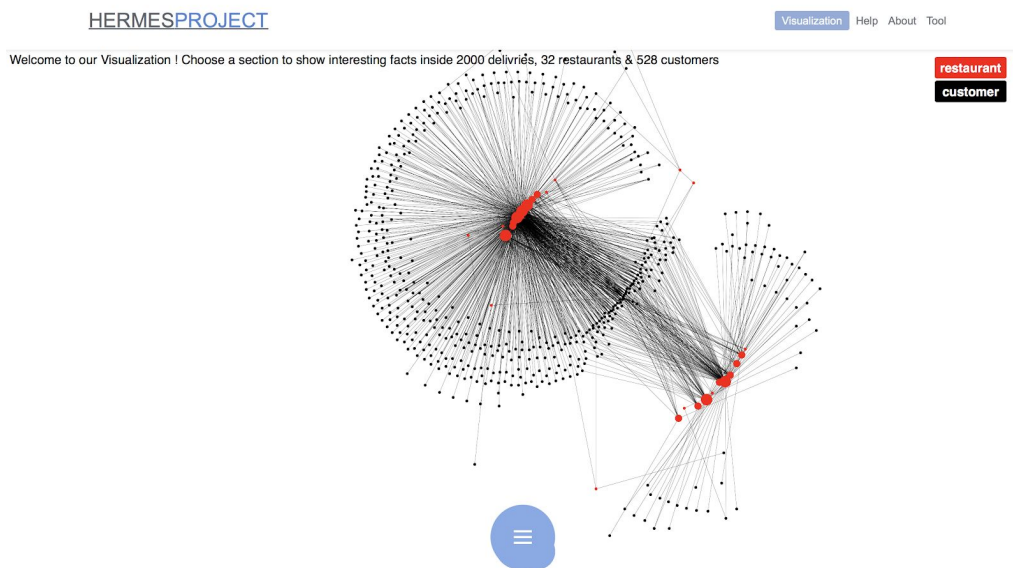


Colorful traces

Visualization of the food deliveries of *Smood*



Process book

Alexandre Colicchio

Mohamed Ndoye



Introduction

In this process book we will describe the development process of our project in the Data Visualization class (COM-480) of the fall 2018 semester.

During the lectures, we learned how to make efficient visualizations using the common tools and languages. From web development to design, a lot of knowledge were given and we saw the do's and don'ts in data visualization.

Overview

Smood is a company specializing in the promotion, sale and delivery of home and office restaurant meals (<https://www.smood.ch/>). Their platform optimizes the ordering and delivery of meals by connecting customers and restaurants equipped with order management terminals. The customers can make an order either on the app or on the website. Founded in 2013, *Smood* is now the leader in food delivery in Switzerland.

Motivation

As a leader of food delivery in Switzerland, we would like to focus in Lausanne and see how *Smood* is operating there. In other words, we would like to explore the deliveries there and gain insights on them. Through our visualization it will be possible to see the details of the deliveries in order to better understand them and to permit the managers of the company to take better decisions.

Target audience

Motivated by the above, our target audience is the managers of *Smood*. We want to give them an useful visualization of their data that will help them understand their operations and so enhance the services all in an easy way.

Concept

In this part we will present the ideas that guide us during the development process. The description of the visualization is established here along with our objectives and what we should obtain at the end. We will cover the description of the dataset, the designs with some sketches, the implementation model, the interaction of the user and finally the differences from the initial ideas will be presented.

Dataset

We have been working with the data deliveries of *Smood*, it contains information on the deliveries that are the pickup and the dropoff locations. It also contains the time when the order was triggered and different values to trace the delivery-man but we didn't use this last parameter.

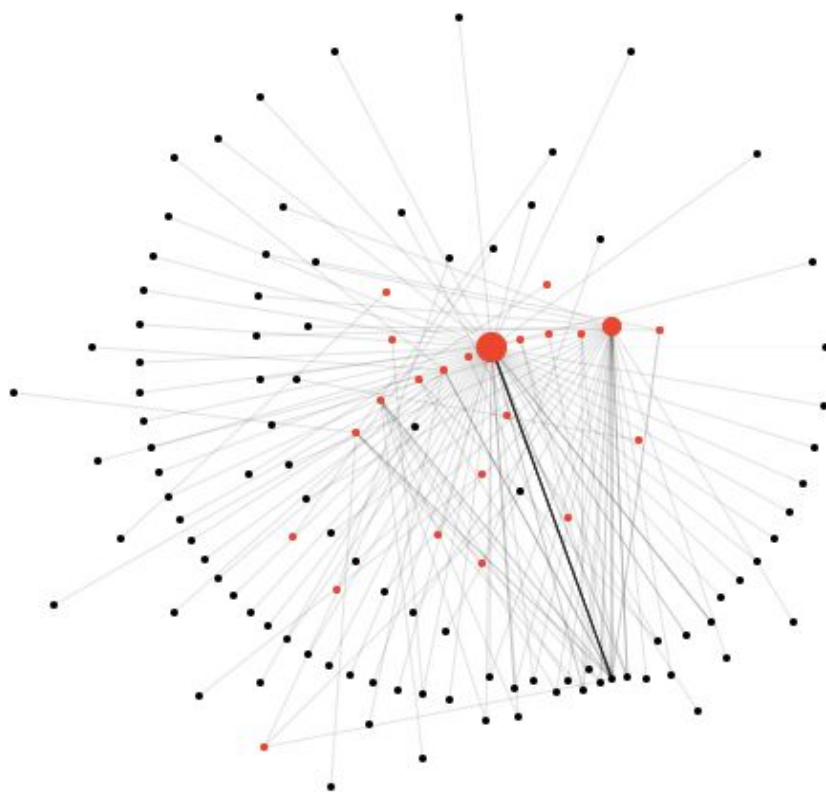
Obtaining the dataset required us to sign an NDA document and because of that we won't give very detailed information on the dataset but will tell the parameters that permit to understand and use the visualization.

We received the data in CSV format and we'll present in the following what we have used :

- pickup_longitude: the longitude where the delivery was engaged
- pickup_latitude: the latitude where the delivery was engaged
- dropoff_longitude: the longitude where the delivery was disengaged
- dropoff_latitude: the latitude where the delivery was disengaged
- time : the time when the order was triggered.

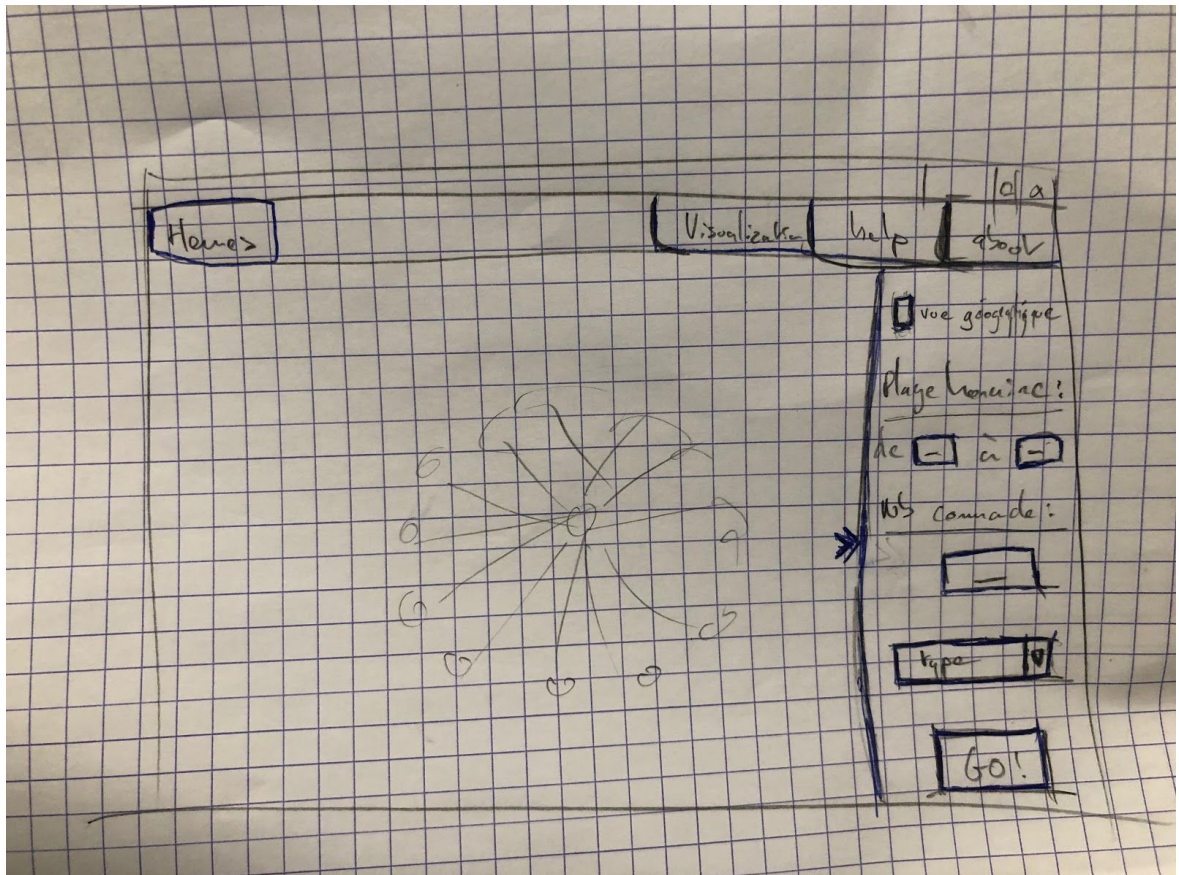
As you can see the dataset is very similar to the well known Kaggle dataset that represents New York City taxi trip duration (<https://www.kaggle.com/c/nyc-taxi-trip-duration/data>).

We processed these information to detect the key restaurants and the key customers, more precisely the restaurants with high demand and the customers that order a lot. We also processed the information to locate the points approximately on a 2D plan.



Designs and interaction

First of all, we decided to put our visualization in a clear and concise website and were inspired by the sketch below.



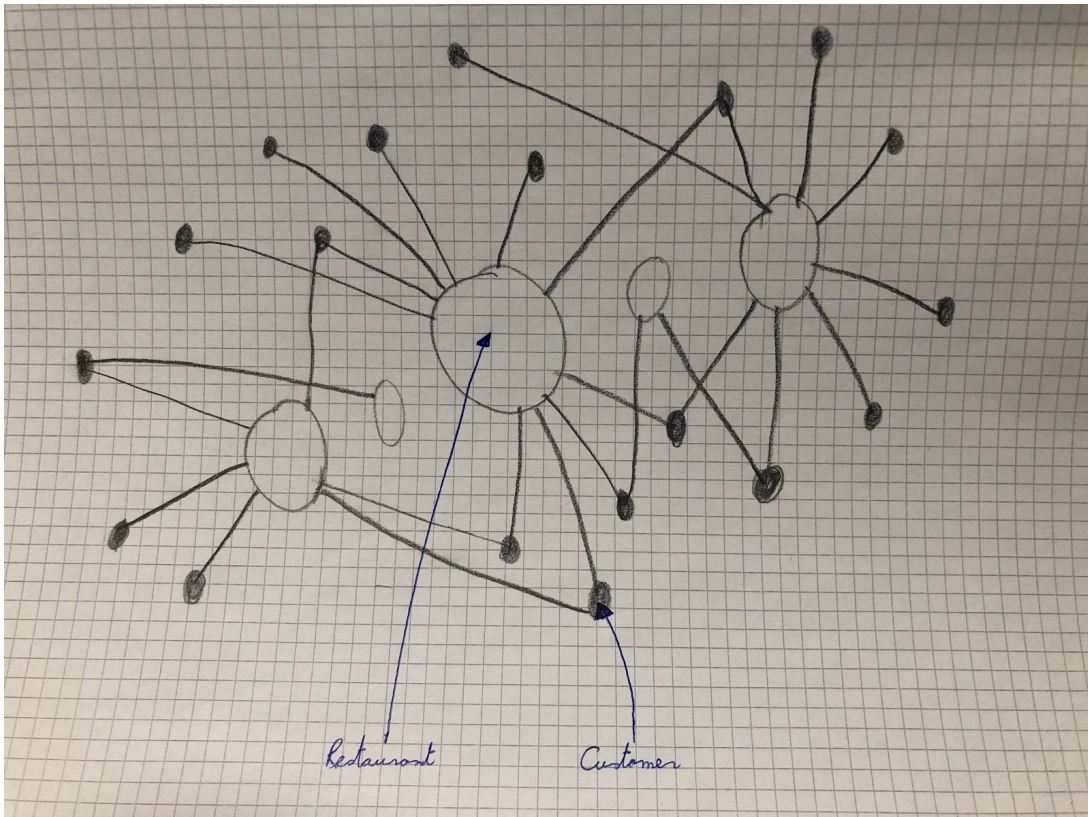
This view appears by clicking on the tool button (upper right) and opening the menu. It allows you to see the different filters and we will explicit their purposes in the interaction section. The goal of this sketch was to establish the framework of the visualization. At the top, there are the main buttons and the logo of the project. We didn't put

a

lot of information here in order to extend the space for the visualization and catch the eye of the user into the viz.

We decided to use graph visualizations, it is the best type of visualization in our case because of the following reasons:

- orders are well represented with an edge between the client and the restaurant, both parties can be represented by nodes
- we can quantify how a restaurant is liked by customers using the size of the node
- finally, different colors can be used to distinguish between the restaurants and the clients



Now we will describe the different interactions.

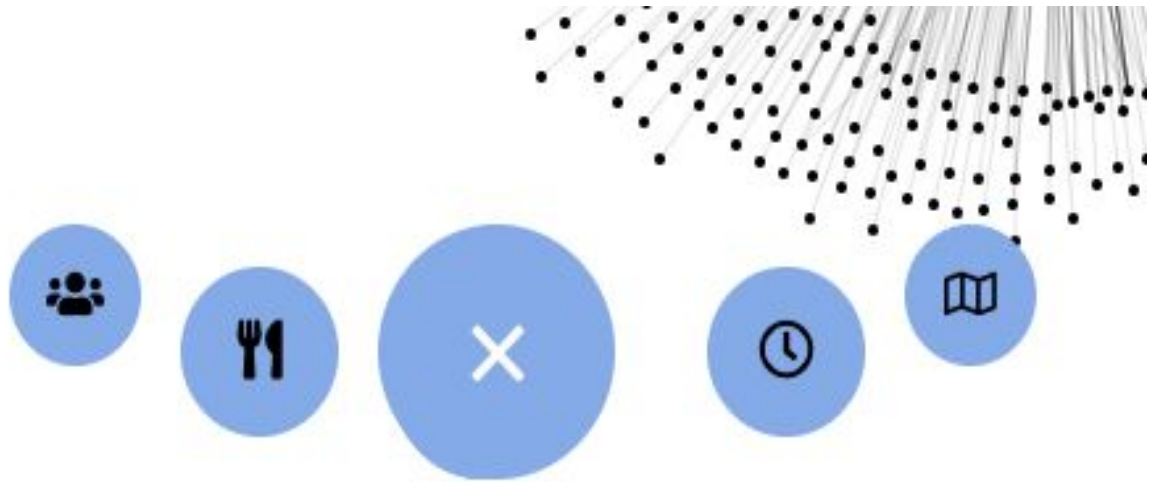
By clicking on a node, either a customer or a restaurant, its number of orders and its address is displayed. The latter is found using the API of Google.

For more precision, it is possible to zoom in, to zoom out and to drag the viz.

By opening the menu as in the picture below, it is possible to select a

topic and display facts along with the corresponding viz on that:

- the first button (leftmost) displays facts on the customers
- the second one displays facts on the restaurants
- the third one displays facts related to the time
- the last one shows facts related to the map

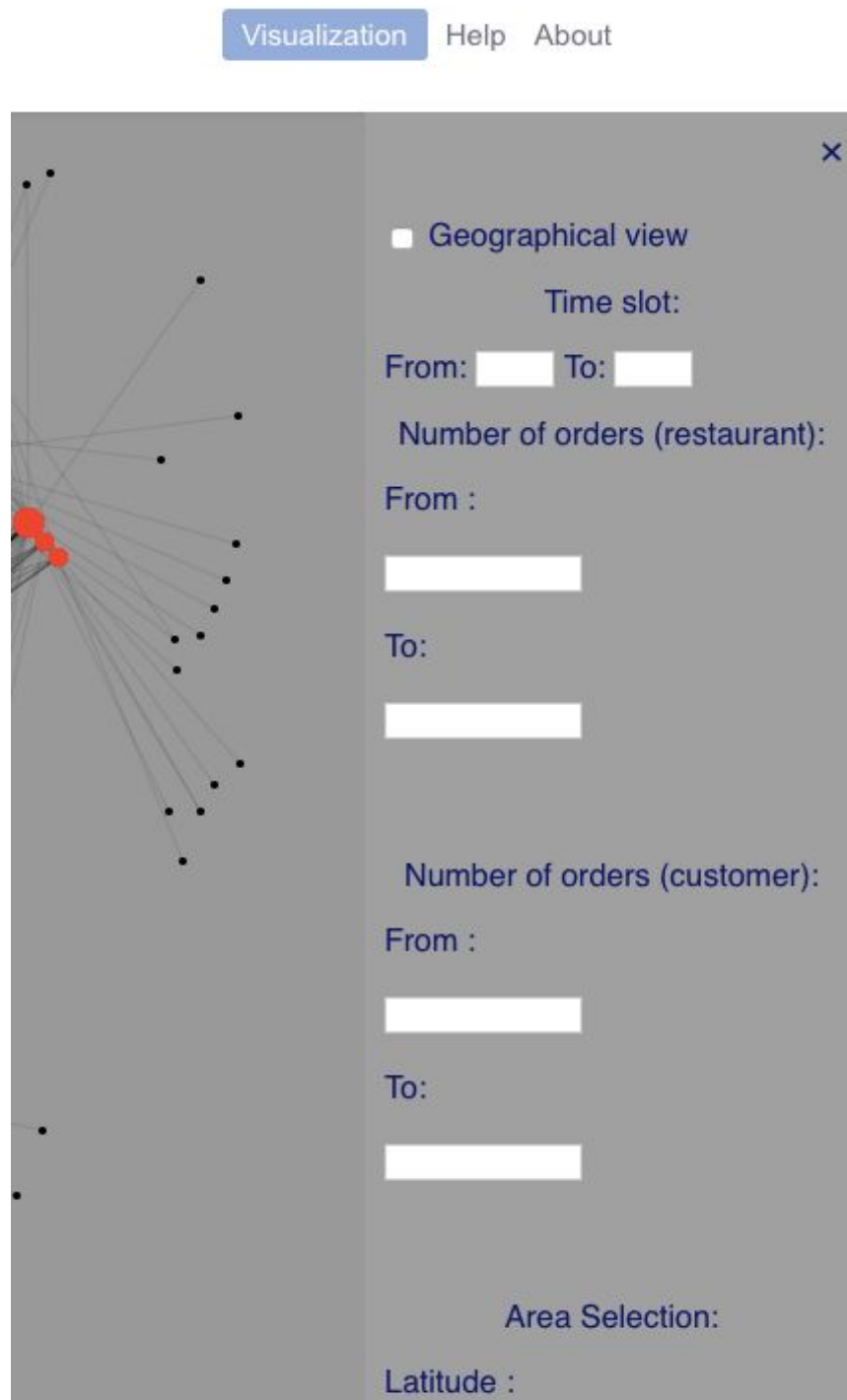


There is also a “tool” mode that allows you to take control of the viz, so of what you want to see, or on what you want to gain insights. In the tool mode, when opening the menu there are five different filters that we describe in the following:

- the graphical view (default) allows you to see the key points
- the geographical view allows you to see the key points in their real positions in the map using their latitude and longitude values
- with the time slot parameter, you can visualize the orders on a given period of the day in order to see when and where the activities are high
- you can also lighten or densify the visualization by filtering by the number of orders you want to see, for restaurants and customers
- the area selection filter allows you the visualize the orders in the selected area

After setting one or several of these parameters, just click on the

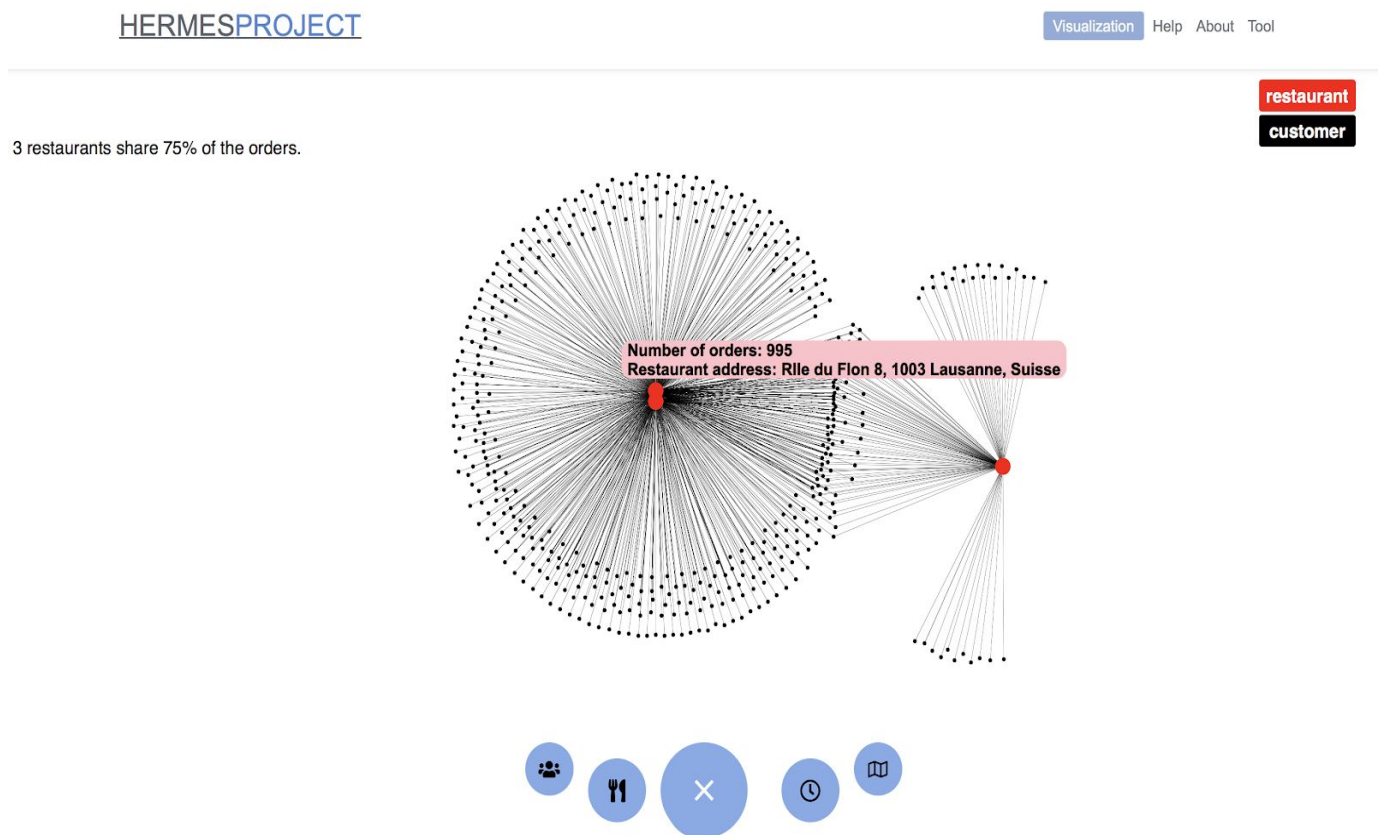
button at the bottom to see the results. It is possible to go back to the initial view with the restart button.



The strength of our work comes from the fact that all our discoveries around the dataset were made using the search tool that we started to create. In other words, this tool can generate an infinity of facts on this data set. This was one of the goals of our work: to provide smood

managers, in addition to a pre-study of their data set, with a tool allowing them to carry out this process themselves.

The visualization follows the same idea: we propose a very abstract visualization, limiting the number of elements on the screen (restaurant, customer and delivery), which allows anyone to quickly understand and extract information.



Differences from the initial proposal

We did deviate a little bit from the initial proposal. We planned a map to give the possibility to the user to trace the itineraries relating to a

restaurant or a delivery man by its corresponding identification number. We also planned to visualize orders by focusing on a specific geographical area or by food type.

Unfortunately, finding information about the different restaurants was impossible (see Slack Smood chat) for the reason that the dataset is old and the restaurants close very quickly.

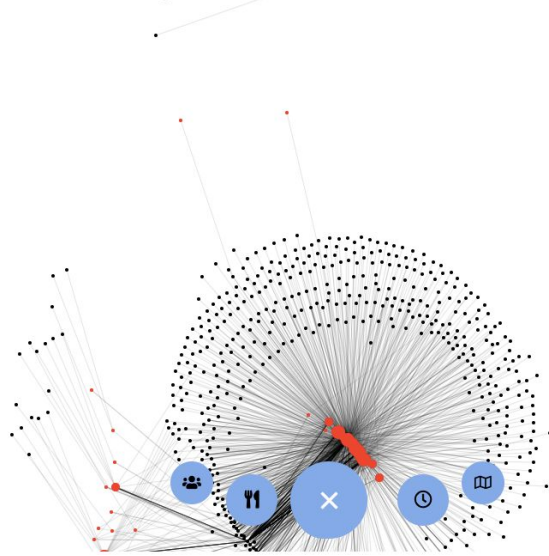
We set up the Google Places API which allowed us to find restaurants in a few meters area around the coordinates provided in the dataset, but for 90% of restaurants, Places API could not find restaurants anymore .

We decided in the end not to put a map for two reasons: the first is that a good number of groups wanted to work on maps, which is in our opinion obvious in relation to the project and that does not bring an "original" visualization. We therefore preferred to propose a more original and simpler "graph" view to extract information, and a geographical view, which takes up the concept of placement in the information space, without overloading the interface with a map. Thus, we can at a glance get the information without getting lost in a map too dense.

In addition, the geographical area was limited here to the city of Lausanne, the map did not give much more information than the street names, which we can get here with a click on the points.

Implementation

In this part, we will describe the essential tools that allowed us to create the different viz that we described above.



We used jQuery for interactions with the visualization menus.

We split the utility functions into `utilitaires.js` and the change of interfaces are implemented in `drawer.js`.

The modification of the display by the filters and the change of the data story is simply done via the modification of the list of restaurants, customers and deliveries.

Finally, the initial import fills these lists with the objects described in `objects.js`.

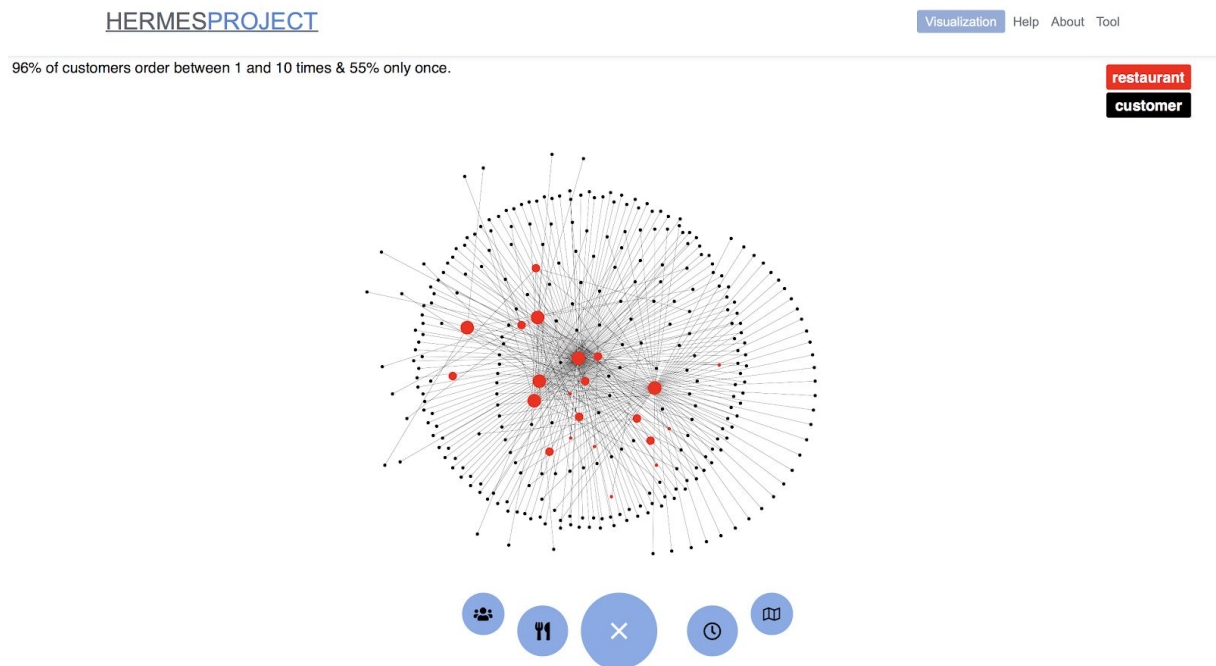
Evaluation

Data insights

With this visualization we gain many insights about *Smood* orders. We will explicit the data story in the following. We find these statistics by using the tool (on the navigation bar) and the different filters that it has.

Starting from the customers:

- 96% of customers ordered between 1 to 10 times and 55% of customers ordered just 1 time
- 20 customers ordered between 10 and 20 times
- 1% of customers ordered between 20 and 30 times
- finally 2 customers ordered more than 30 times, in 13 different restaurants.

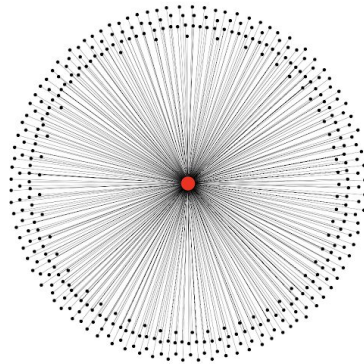


Let's take a look at the restaurants:

- the best restaurant cumulates to itself 995 orders, around the half of the total orders
- 50% of restaurants accumulate between 1 to 10 orders
- 3 restaurants share 75% of the orders

The best restaurant has 995 orders!

restaurant
customer

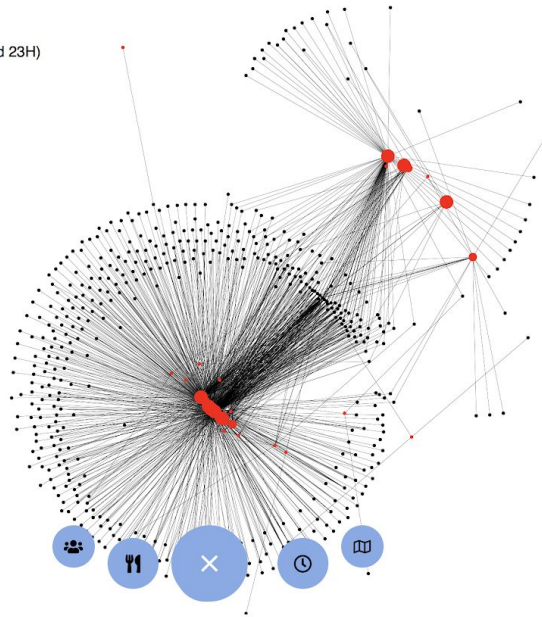


Regarding the time:

- 3% of the orders are before noon whereas 97% of them are afternoon
- there are only 46 orders between 12h and 14h
- 89% of food are delivered on evening, between 17h and 23h

89% of meals are delivered in the evening (between 17H and 23H)

restaurant
customer

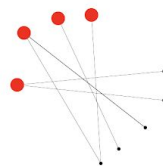


Finally let's explore the map:

- a quarter of the order are made in the center of Lausanne
- half of the clients are located at 2 km within the city center
- there are 5 clients on EPFL campus

There are 5 clients on the EPFL campus ! :)

restaurant
customer



Future work

More detailed insights on the dataset could be extracted. For example we could add a way to find the food type that each restaurant offers and distinguish the restaurants between them. Unfortunately, we were not able to do it because with the values of the position of the restaurants given we couldn't find any restaurant at the given position. So it was impossible to detect the food type they offer. As a future work, it would be great to analyze that and find a solution. We have been able to implement consistent features with the data set in our possession. We tried to add additional datas, especially through the Google APIs. However, a more consistent data set would have undoubtedly allowed us to set up a lot of additional functionality and views.

Peer assessment

Despite the fact that we were two instead of three, we managed to do a good job. We were both on point during the meetings and contribute productively to the discussion and work. Each one was able to give his ideas and find solutions whenever disagreements occurred.